



Associations between genetic liabilities to smoking behavior and schizophrenia symptoms in patients with a psychotic disorder, their siblings and healthy controls

Bochao Danae Lin^{a,b,c}, Jentien M. Vermeulen^d, K. Bolhuis^e, Xiao Chang^f, Frederike Schirmbeck^{a,g}, Kristel R van Eijk^{a,b}, Genetic Risk and Outcome of Psychosis (GROUP) investigators[#], Sinan Guloksuz^{c,h}, Matthijs Blankers^{g,i}, W. van den Brink^d, Lieuwe de Haan^{c,g}, Jurjen J. Luykx^{a,b,j,*}

^a Department of Psychiatry, University Medical Center Utrecht, Utrecht University, Brain Center Rudolf Magnus, Utrecht, The Netherlands

^b Department of Translational Neuroscience, University Medical Center Utrecht, Utrecht University, Brain Center Rudolf Magnus, Utrecht, The Netherlands

^c Department of Psychiatry and Neuropsychology, School for Mental Health and Neuroscience, Maastricht University Medical Centre, Maastricht, The Netherlands

^d Department of Psychiatry, Amsterdam UMC location AMC, University of Amsterdam, Amsterdam, The Netherlands

^e Department of Child and Adolescent Psychiatry, Erasmus Medical Center-Sophia Children's Hospital, Rotterdam, The Netherlands

^f Institute of Science and Technology for Brain-Inspired Intelligence, Fudan University, Shanghai, 200433, China

^g Arkin Mental Health Care, Amsterdam, The Netherlands

^h Department of Psychiatry, Yale University School of Medicine, New Haven, CT

ⁱ Trimbos institute – The Netherlands institute of Mental Health and Addiction, Utrecht, The Netherlands

^j GGNet Mental Health, Warnsveld, The Netherlands

ARTICLE INFO

Keywords:

Polygenic
Smoking
Schizophrenia
Psychosis
Positive symptoms
Negative symptoms

ABSTRACT

It is unknown how smoking behavior polygenic scores (PRS) relate to psychosis and psychotic symptoms. To elucidate this, genotype and phenotype data were collected from patients with schizophrenia, their unaffected siblings, and healthy controls in a six-year follow-up prospective cohort study. Associations between smoking behaviors, PRS and schizophrenia symptoms were explored using linear mixed-effect models. The mean number of cigarettes smoked per day were 18 for patients, 13 for siblings and 12 for controls. In the overall sample, PRSs-smoking initiation (i.e., ever smoking as a binary phenotype, PRS-SI) were positively associated with positive symptoms, negative symptoms, and depressive symptoms, whereas PRSs-AI (age at regular smoking initiation) were negatively associated with all symptom dimensions, with similar effect sizes. When considering groups separately, PRS were only associated with psychotic symptoms in siblings and controls. In conclusion, unaffected siblings show smoking behaviors at an intermediate level between patients and healthy controls. Additionally, PRS-SI and PRS-AI are associated with all symptom dimensions only in unaffected siblings and healthy controls, possibly owing to the dominant role of other (genetic) risk factors in patients. Future studies may examine mechanisms via which genetic risk for smoking affects mental health symptoms.

* Corresponding author.

E-mail address: j.luykx@maastrichtuniversity.nl (J.J. Luykx).

[#] GROUP investigators are: Behrooz Z. Alizadeh^{a,b}, Agna A. Bartels-Velthuis^a, Therese van Amelsvoort^g, Richard Bruggeman^{a,c}, Wiepke Cahn^{d,k}, Lieuwe de Haan^{f,m}, Bart P.F. Rutten^g, Sinan Guloksuz^g, Jurjen J. Luykx^{d,e,i}, Jim van Os^{d,j}, Claudia J.P. Simons^{g,i}, Ruud van Winkel^{g,ha} University of Groningen, University Medical Center Groningen, University Center for Psychiatry, Rob Giel Research center, Groningen, The Netherlands; ^bUniversity Medical Center Groningen, Department of Epidemiology, Groningen, The Netherlands; ^cUniversity of Groningen, Department of Clinical and Developmental Neuropsychology, Groningen, The Netherlands; ^dUniversity Medical Center Utrecht, Department of Psychiatry, Brain Centre Rudolf Magnus, Utrecht University, Utrecht, The Netherlands; ^eUniversity Medical Center Utrecht, Department of Translational Neuroscience, Brain Center Rudolf Magnus, Utrecht, The Netherlands; ^fAmsterdam UMC, University of Amsterdam, Department of Psychiatry, Amsterdam, The Netherlands; ^gMaastricht University Medical Center, Department of Psychiatry and Neuropsychology, School for Mental Health and Neuroscience, Maastricht, The Netherlands; ^hKU Leuven, Department of Neuroscience, Research Group Psychiatry, Leuven, Belgium; ⁱGGzE Institute for Mental Health Care, Eindhoven, The Netherlands; ^jKing's College London, King's Health Partners, Department of Psychosis Studies, Institute of Psychiatry, London, United Kingdom; ^kAltrecht, General Mental Health Care, Utrecht, The Netherlands; ^lGGNet Mental Health, Warnsveld, The Netherlands; ^mArkin, Institute for Mental Health, Amsterdam, The Netherlands.

<https://doi.org/10.1016/j.psychres.2023.115143>

Received 15 December 2022; Received in revised form 21 February 2023; Accepted 1 March 2023

Available online 6 March 2023

0165-1781/© 2023 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Smoking is probably the single most unhealthy human behavior (Revicki, Sobal, & DeForge, 1991). Although the prevalence of tobacco smoking in the general population has decreased in the past decades, smoking rates among patients with schizophrenia are still very high (up to 70%) (Lasser et al., 2000). The odds that patients with schizophrenia smoke, is more than three times higher than for people from the worldwide general population (de Leon & Diaz, 2005; Zeng et al., 2020). Such high rates of smoking lead to tobacco-related diseases that substantially increase early mortality rates in patients with schizophrenia. For example, the standardized mortality rate for chronic obstructive pulmonary disease-related deaths is 9.9, CI 9.6-10.2 (Olsson, Gerhard, Huang, Crystal, & Stroup, 2015).

Increased rates of smoking are observed in first-degree relatives of individuals with psychotic disorders, albeit to a lesser degree than in patients (Lyons et al., 2002; Vermeulen et al., 2018), giving rise to the possibility of genetic overlap between psychosis and smoking behaviors, which is supported by findings from hypothesis-generating genetic studies: in a genome-wide association study (GWAS), single-nucleotide polymorphisms (SNPs) in the human *CHRNA5-A3-B4* cluster – a promising candidate region for smoking behaviors – were significantly associated with schizophrenia (Schizophrenia Working Group of the Psychiatric Genomics, 2014; Trubetskoy et al., 2022). This gene cluster, located at chromosome 15, encodes an nAChR subunit that is associated with an increased risk for heavy smoking but not with smoking initiation (Liu et al., 2019). Moreover, polygenic risk scores (PRSs), weighted sums of trait-associated alleles, of the nicotine metabolite cotinine are significantly associated with PRSs for schizophrenia (Chen et al., 2016). In addition, PRSs for age at smoking initiation were significantly associated with schizophrenia in a Japanese sample (Ohi et al., 2020). Finally, positive genetic correlations ($rg = 0.19$; $p = 0.037$) have been established between schizophrenia and age at onset of smoking and rate of cigarettes smoked ($rg = 0.14$; $p = 0.049$) using LD regression score (<http://ldsc.broadinstitute.org/>).

To our knowledge, these initial observations on the possible shared genetic susceptibility between schizophrenia symptom clusters and smoking behaviors have not been substantiated by family studies examining polygenic liabilities. Findings from such studies can provide further insight into the role of smoking in the causal chain leading to psychosis and its subclinical course of psychotic symptoms, which is not fully understood (Quigley & MacCabe, 2019). Healthy (unaffected) siblings of patients with a psychotic disorder share environmental and half of the genetic risks of the patients and have genetic liabilities to psychosis in between patients (i.e., fully affected) and healthy controls, and are free of illness-specific confounding factors. Including unaffected siblings of patients with a psychotic disorder can thus contribute to disentangle the associations between genetic underpinnings of smoking behaviors and psychosis symptom clusters.

Here, we therefore investigated 1) baseline difference in smoking behavior between patients with a psychotic disorder, unaffected siblings and healthy controls; 2) associations of smoking behavior PRSs with psychotic disorder case-control status; and 3) associations of smoking behavior PRSs with schizophrenia symptom clusters. To that end, we generated PRSs for age at onset of smoking, age at first regular smoking, number of cigarettes smoked per day, and schizophrenia. We then evaluated associations with case-control status and psychosis symptom clusters in a longitudinal cohort of patients with non-affective psychosis, unaffected siblings of patients with psychosis, and healthy controls. Based on the shared-vulnerability hypothesis, we expected higher genetic risk for smoking behaviors to be associated with higher symptom levels in patients and siblings due to shared genetic and environmental factors, and to a lesser extent in healthy controls.

2. Methods

2.1. Study population and study design

This study was performed within the naturalistic, multi-center cohort study of the Genetic Risk and Outcome of Psychosis (GROUP) study (Korver et al., 2012). The full sample consisted of 3,684 participants: 1,119 patients with a diagnosis within the non-affective psychotic spectrum, 1,059 unaffected siblings, 920 parents of patients with psychotic disorders, and 586 unrelated healthy controls. Study design, power calculations, recruitment procedure and baseline characteristics of participants have been described in detail previously (Korver et al., 2012). In short, patients aged between 16 and 50 years, and diagnosed with non-affective psychosis according to the Diagnostic and Statistical Manual of Mental Disorders Fourth Edition (DSM-IV) (*Diagnostic and Statistical Manual of Mental Disorders: DSM-IV-TR, 2000*), were recruited by clinicians from four university medical centers and 36 associated mental health care facilities in representative geographical areas in the Netherlands and Belgium between 2004 and 2014. The age range was chosen to allow for inclusion of early-onset cases as well as long-term follow-up (older age would have incurred higher risks of mortality and thus loss to follow-up). Siblings and controls were included if not affected by a psychotic disorder. All patients, unaffected siblings and controls took part in the baseline assessment (T0) and were invited for follow-up assessments three (T1) and six (T2) years after inclusion (and the baseline visit). Patients were included in the analyses when genetic data and at least one of the outcomes of interest was available.

This study was approved by the Medical Ethics Committee of the University Medical Center of Utrecht. Written informed consent was obtained before inclusion.

2.2. Genetic data and quality control steps

Genotype data for 2,812 individuals were generated on a customized Illumina, IPMCN array with 570,038 SNPs. This chip contains ~250k common SNPs, 250K rare, exomic, non-synonymous SNPs [minor allele frequency (MAF) < 1%], and ~50K psychiatric-related variants. Quality control (QC) procedures were performed using PLINK v1.9 (Purcell et al., 2007) as previously reported (Pazoki et al., 2020); detailed quality control steps are shown in the **Supplementary method, p.4 of the Supplement**. The quality-controlled SNPs were imputed on the Michigan server (Das et al., 2016) using the HRC r1.1 2016 reference panel with European samples after phasing with Eagle v2.3. Post-imputation QC involved removing SNPs with an estimated r^2 (Rsq) info score < 0.3, with a MAF < 0.01, SNPs that had a discordant MAF (MAF difference > 0.15) compared to the reference panel, and strand ambiguous AT/CG SNPs and multi-allelic SNPs, leaving a total of 2505 subjects and 14,132,467 SNPs for final analyses. We excluded the patients' parents ($n=700$) from the sample because of lacking phenotypic information (smoking behavior and CAPE score), resulting in 1805 individuals for analysis, including $n=706$ patients, $n=731$ siblings and $n=368$ healthy controls.

2.3. Polygenic risk score (PRS) calculations

Smoking behavior PRSs in our study population were calculated from GSCAN (consortium GWAS & Sequencing Consortium of Alcohol and Nicotine use): (i) a binary phenotype indicating whether an individual had ever initiated smoking regularly (SI), (ii) age of regular smoking initiation (AI), and (iii) number of cigarettes smoked daily (CPD) (Liu et al., 2019). The exact definition of 'regular' ranged across cohorts from a cumulative number of 100 cigarettes ever smoked and daily smoking for a month to a positive answer to the question: 'Have you ever smoked regularly?' (Khouja, Wootton, Taylor, Davey Smith, & Munafò, 2021). In addition, PRS-schizophrenia were generated based on the latest GWAS of schizophrenia excluding the current study

population (Trubetsky et al., 2022). Odds ratios for autosomal SNPs reported in the schizophrenia summary statistics were log-converted to beta values. Overlapping SNPs between each training dataset (smoking behavior GWASs), 1000 reference Genome (reference dataset), and our GROUP sample dataset (target dataset) were selected. Then the following SNPs were excluded: 1) insertions, deletions and ambiguous SNPs; 2) SNPs with minor allele frequency (MAF) < 0.01; and 3) SNPs with imputation quality (R^2) < 0.8 in both training dataset and target dataset; and 4) SNPs located in 20 complex-LD regions (Price et al., 2008) (see **Supplementary Table S1**), such as *MHC* (major histocompatibility complex) (Price et al., 2008). Overlapping SNPs were clumped in two rounds using PLINK v1.9; round 1 with the default parameters (physical distance threshold 250kb and LD threshold (R^2) < 0.5) and round 2 with a physical distance threshold of 5,000kb and LD threshold (R^2) < 0.2; resulting in 269,813 SNPs for PRSs calculation. PRSs were calculated using PRSice2 (Choi, Mak, & O'Reilly, 2020) with default settings for 12 GWAS p-value thresholds (p_t): 5×10^{-8} , 5×10^{-7} , 5×10^{-6} , 5×10^{-5} , 5×10^{-4} , 5×10^{-3} , 0.05, 0.1, 0.2, 0.3, 0.4 and 0.5.

2.4. Levels of (sub)clinical symptoms clusters

Psychosis-related symptoms in patients, siblings and controls were assessed with the self-rated Community Assessment of Psychic Experiences (CAPE; (Mossaheb et al., 2012)). The CAPE is a self-report questionnaire to assess schizophrenia symptom clusters. At baseline (T0), the CAPE assessed lifetime symptoms; at follow-up T1 and T2 CAPE assessed symptoms over the past three years. Each of the items is rated in terms of frequency on a scale from 0 (absent) to 3 (severe). Mean scores of all items (0-3) representing positive symptoms (20 items), negative symptoms (14 items) and depressive symptoms (8 items) were calculated for patients, siblings, and healthy controls.

2.5. Statistical analyses

Demographic variables, age of smoking initiation and numbers of cigarettes smoked per day, were compared across groups (patients, siblings, and healthy controls) using ANOVA and t-tests (p-value threshold for significance < 0.05).

To study the associations of smoking behavior PRSs (PRS-SI, PRS-AI and PRS-CPD) with (1) psychotic disorder-case control status; and (2) psychotic symptom levels, linear mixed-effects regression analyses were conducted using R version 3.6.2 (R Core Team, 2016) and the lme4 package (Bates, Machler, Bolker, & Walker, 2015).

For 1), we performed a logistic regression using cross-sectional data (time point= T0; N= 1084, including 706 patients and 368 healthy controls) to detect associations between psychotic disorder case-control status (Y) with PRS-smoking behaviors, the first 3 genetic principal components (PCs), age, and sex as fixed effects as follows:

$$Y = \beta_0 + \beta_1 * PRS + \beta_2 * Age + \beta_3 * Sex + \beta_4 * PC1 + \beta_5 * PC2 + \beta_6 * PC3 + \varepsilon$$

For 2), we first performed these analyses in the entire study population and then in patients, sibling, and controls separately. The first 3 genetic PCs, time, age, sex, and sibling effects (0=control, 1=siblings and 2=patients) as fixed effects were added to the association models as covariates. In addition, we added intercepts for subjects and by-subject random slopes for the effects of time and family structure as random effects for association tests in the entire cohort. The equation for the entire study population was:

$$Y_{1..3} = \beta_0 + \beta_1 * PRS + \beta_2 * Age + \beta_3 * Sex + \beta_4 * PC1 + \beta_5 * PC2 + \beta_6 * PC3 + \beta_7 * sibling + \beta_8 * SCZ + Z * FamilyID + Z * Time + \varepsilon$$

In the subgroups (patients, relatives, and controls) the equation was:

$$Y_{1..3} = \beta_0 + \beta_1 * PRS + \beta_2 * Age + \beta_3 * Sex + \beta_4 * PC1 + \beta_5 * PC2 + \beta_6 * PC3 + Z * FamilyID + Z * Time + \varepsilon$$

All variables were added *en bloc* and models were fitted with restricted maximum likelihood models (REML). To report how much variance is explained by the risk score itself, the delta r^2 for fixed effects were calculated in linear mixed models with PRSs and without PRSs. P-values were calculated with the Kenward-Roger approach (Kuznetsova, Brockhoff, & Christensen, 2017). As the PRSs were calculated in 12 p-values thresholds, the association p-values in all PRS analyses were adjusted for false discovery rate (FDR) of these 12 PRSs (p-value threshold for significance < 0.05) (Benjamini & Hochberg, 1995). To test whether the PRS-smoking associations would still remain when accounting for genetic liability to schizophrenia, we added PRS-SCZ ($p_t=0.05$) as a covariate to the abovementioned models. Following suggestions put forward during the peer review process, we also checked whether the results for patients on olanzapine, clozapine or both were similar to the results for the entire group of patients by restricting our analyses to such patients as a subgroup.

We then split samples by PRS ($p_t=0.05$) tertiles into low, middle and high PRS groups. Symptoms were then compared between low, middle and high PRS-SI, PRS-AI and PRS-CPD groups in the entire cohort using Wilcoxon rank-test (p-value threshold for significance < 0.05). Finally, we tested interaction effects of patient/siblings control status (group status) with PRSs ($p_t=0.05$) to check whether the PRS association results remained. GROUP database release 5.0 was used for all analyses.

3. Results

3.1. Smoking behavior differences between cases, unaffected siblings and health controls

In total, genotype data of 706 patients with non-affective psychosis, 731 unaffected siblings and 368 healthy controls (total N = 1805) were available after we executed the genetic quality control steps, as described above.

The percentage of smokers in the different groups was 64% (n=454) of the patients, 38% (n=277) of the siblings and 23% (n=85) of the controls.

The mean (SD) age of smoking initiation was 15.4 (3.6) years for patients, 15.7 (2.9) years for siblings and 16.1 (3.6) years for controls (ANOVA test: $F = 14.9$, $p = 4.4 \times 10^{-7}$, Fig. 1A).

The mean (SD) number of cigarettes smoked per day for regular smokers was 17.8 (9.12) for patients, 12.6 (7.98) for siblings and 11.5 (7.32) for controls (ANOVA test, $F=61.8$, $p = 8.14 \times 10^{-153}$, Fig. 1B).

CAPE scores in patients showed a mean (SD) item score of 0.67 (0.49) for positive symptoms, 1.01 (0.53) for negative symptoms and 0.97 (0.56) for the depressive subscale. These and other characteristics for the three groups (patients, unaffected siblings, healthy controls) are reported in Table 1.

For the association analyses with schizophrenia symptoms, varying by the completeness of CAPE subscale scores, we used data on 686-687 patients with 1601-1602 observations, 575-579 siblings with 966-971 observations and 298 controls with 538 observations. The specific sample sizes for association tests are listed in each respective table (**Supplementary Tables S2-5**).

Association analyses between smoking behavior PRSs, psychotic disorder case-control status and symptom levels

The most significant association was found between psychotic disorder case-control status and PRS-CPD at $p_t = 5 \times 10^{-3}$ (beta= 0.105, SE=0.053, $P_{FDR}=0.003$; **Supplementary Figure S1**). However, when adding PRS-SCZ to the model as a covariate (fixed), the association between PRS-CPD and psychotic disorder case-control status was not significant anymore.

PRSs-SI were positively associated with positive ($p_t 5 \times 10^{-5}$ -0.5,

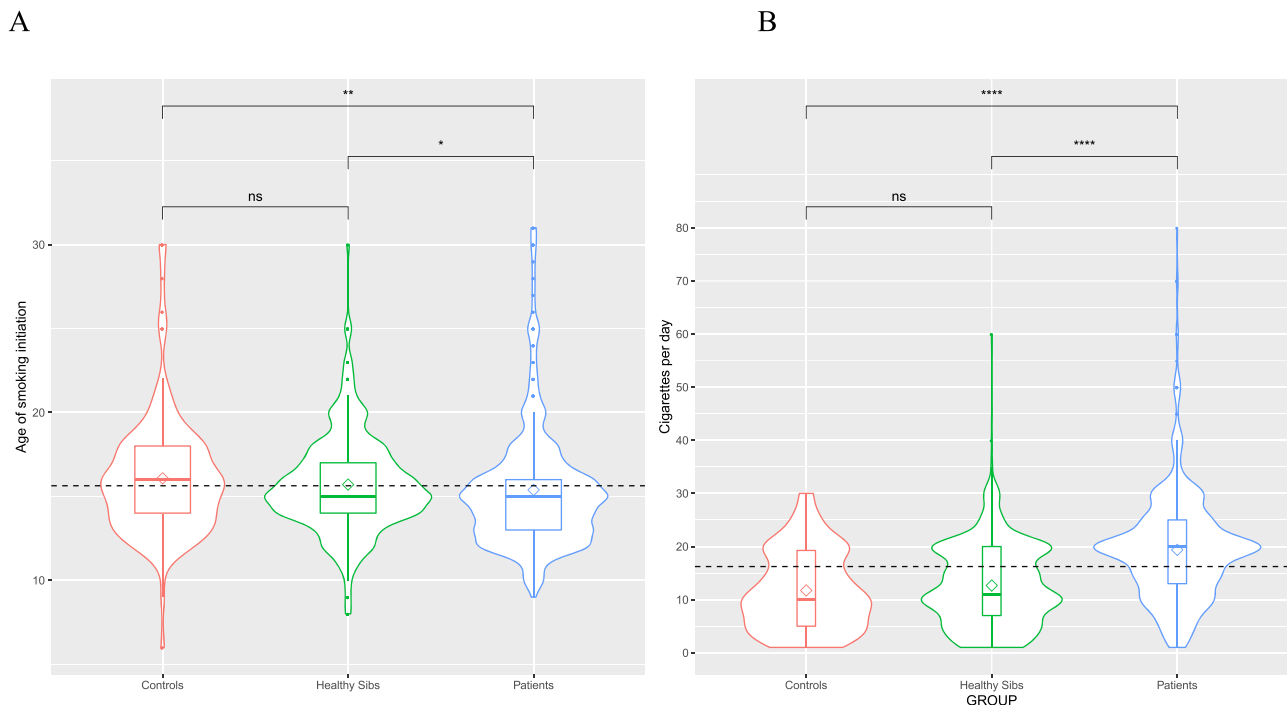


Fig. 1. Smoking behavior distribution in patients with a psychotic disorder, siblings and healthy controls. Figure 1A. Violin and box plots of age of smoking initiation in healthy controls, siblings and patients (years). Figure 1B. Violin and box plots of number of cigarettes smoked per day in healthy controls, siblings and patients. The dashed line denotes the mean age of smoking initiation or cigarettes per day in all participants, respectively. The diamonds denote the mean age of smoking initiation and cigarettes per day in each group, respectively. Mean comparisons between pairs were examined using T-tests: ns: $p > 0.05$, *: $p < 0.05$, ** $p < 5.0 \times 10^{-3}$, ****: $p < 5.0 \times 10^{-5}$.

optimal $p_t = 0.05$, $\beta = 0.023$, $SE = 0.01$, $R^2 = 0.003$, $P_{FDR} = 0.045$; Fig. 2A), negative ($p_t = 5 \times 10^{-6}$ -0.5, optimal $p_t = 5 \times 10^{-4}$, $\beta = 0.135$, $SE = 0.036$, $R^2 = 0.007$, $P_{FDR} = 0.003$; Fig. 2B) and depressive symptoms ($p_t = 5 \times 10^{-5}$ -0.5, optimal $p_t = 5 \times 10^{-4}$, $\beta = 0.136$, $SE = 0.036$, $R^2 = 0.007$, $P_{FDR} = 0.002$; Figure 2C; Table S2A).

In line with this finding, PRSs-AI were negatively associated with positive (optimal $p_t = 0.05$, $\beta = -0.014$, $SE = 0.006$, $R^2 = 0.002$, $P_{FDR} = 0.047$), negative (optimal $p_t = 0.4$, $\beta = -0.017$, $SE = 0.005$, $R^2 = 0.005$, $P_{FDR} = 0.007$) and depressive symptoms (optimal $p_t = 0.4$, $\beta = -0.014$, $SE = 0.005$, $R^2 = 0.004$, $P_{FDR} = 0.035$; Suppl. Table S2A; Suppl. Figure S2). PRSs-CPD were not significantly associated with positive and depressive symptoms, but only positively associated with negative symptoms (optimal $p_t = 0.05$, $\beta = 0.028$, $SE = 0.008$, $R^2 = 0.005$, $P_{FDR} = 0.012$; Suppl. Table S2A; Suppl. Figure S3). When adding PRS-SCZ to the model as a covariate (fixed), the results did not change, except for the associations of PRS-CPD with depressive symptoms (Suppl. Table S 2B).

Association analyses between smoking behavior PRSs and psychotic symptoms in patients, siblings and controls separately

Strongest associations between PRS-SI and psychotic symptoms were found in healthy controls, while no significant associations were found in patients (with similar results in patients using clozapine, olanzapine, or both); siblings showed intermediate strengths of associations (Suppl. Table S3A; Fig. 3).

In healthy controls, PRS-SI were associated with positive (optimal $p_t = 0.05$, $\beta = 0.036$, $SE = 0.009$, $R^2 = 0.031$, $P_{FDR} = 0.002$; Fig. 3A) and negative symptoms (optimal $p_t = 0.05$, $\beta = 0.052$, $SE = 0.019$, $R^2 = 0.018$, $P_{FDR} = 0.028$; Fig. 3B). In siblings, we found an association between PRS-SI and positive symptoms (optimal $p_t = 0.05$, $\beta = 0.02$, $SE = 0.008$, $R^2 = 0.006$, $P_{FDR} = 0.04$; Fig. 3A). All these results remained significant when adding PRS-SCZ as a covariate (Suppl. Table S3 B).

For PRS-AI, the association was only significant in siblings (Suppl. Table S4A; Suppl. Figure S4): PRS-AI were associated with negative symptoms (optimal $p_t = 0.5$, $\beta = -0.020$, $SE = 0.006$, $R^2 = 0.011$,

$P_{FDR} = 0.014$) and depressive symptoms (optimal $p_t = 0.5$, $\beta = -0.018$, $SE = 0.006$, $R^2 = 0.008$, $P_{FDR} = 0.025$) in siblings. As PRS-AI captures age at regular smoking initiation, the direction of effect was in line with the other findings. All these results remained significant when adding PRS-SCZ as a covariate (Suppl. Table S4B).

For PRS-CPD, we only found significant associations in siblings (Suppl. Table S5 A; Suppl. Figure S5): PRS-CPD were positively associated with negative symptoms ($p_t = 0.05$, $\beta = 0.037$, $SE = 0.011$, $R^2 = 0.013$, $P_{FDR} = 0.007$). All these results remained significant when adding PRS-SCZ as a covariate (Table S5 B).

When adding PRS*patient/siblings/controls status to the model (Suppl. Table S6), the main associations of PRS-SI with positive, negative, and depressive symptoms remained ($P < 0.05$), with no significant interactions detected. Correlations between PRS are reported in Suppl. Table S7.

4. Discussion

Here, we found that unaffected siblings of patients with psychotic disorders show smoking behaviors at an intermediate level between patients and healthy controls. Furthermore, we demonstrate that polygenic risk scores for smoking initiation and age of smoking initiation (PRS-SI and PRS-AI) are associated with schizophrenia symptom clusters only in unaffected siblings and healthy controls. In patients, no associations were found between genetic liabilities for smoking phenotypes and schizophrenia symptom levels. These findings suggest that genetic liabilities to smoking behaviors are differentially related to schizophrenia symptom clusters.

Our results suggest that many SNPs related to smoking initiation with small individual effect sizes together contribute 3.1% and 1.9% to the variance of positive, negative and depressive symptom levels in individuals without illness-related confounders. Possible explanations for this finding could include a shared underlying biological pathway such as the cholinergic receptor or horizontal pleiotropy (Quigley &

Table 1
Study sample demographics at baseline.

	Patients (N=706) Mean (SD)	Siblings (N=731) Mean (SD)	Healthy Controls (N=368) Mean (SD)	P-value
Age (years)	27.77 (8.20)	27.40 (8.11)	30.5 (10.72)	3.58×10^{-12}
CAPE (0-3) ¹				
• Positive symptoms ²	0.67(0.49)	0.21(0.20)	0.20(0.18)	1.12×10^{-268}
• Negative symptoms ³	1.01(0.53)	0.56(0.39)	0.51(0.34)	3.72×10^{-174}
• Depressive symptoms ⁴	0.97(0.56)	0.63(0.40)	0.63(0.37)	4.20×10^{-101}
Smoking initiation age (years) ⁵	15.38(3.60)	15.70(2.86)	16.07(3.55)	0.0014
CPD in smokers ⁶	18.16(8.84)	12.92(7.85)	11.84(7.36)	8.12×10^{-153}
Sex(male)	N (%) 539(76)	N (%) 335(46)	N (%) 161(44)	Chi-square test among 3 groups < 2×10^{-16}
Smoking status (yes) ⁷	445(64)	271(37)	85(23)	< 2×10^{-16}
Cannabis, urine positive ⁸	73(10)	38(6)	13(4)	1.2×10^{-6}
Antipsychotic drug use (yes) ⁹	474(90)			

¹ A mean score of items representing positive symptoms (for at least 14 out of 20 items), negative symptoms (for at least 9 out of 14 items), and depressive symptoms (for at least 5 out of 8 items) was calculated. ²Baseline data were missing for 151 patients, 82 siblings and 23 controls. ³Baseline data were missing for 152 patients, 82 siblings and 23 controls. ⁴Baseline data were missing for 148 patients, 82 siblings and 21 controls. ⁵Age at first smoking cigarettes was only available in baseline data; baseline data were missing for 378 patients, 424 siblings and 234 controls. ⁶Baseline data were missing for 3 patients, 2 siblings and 3 controls. ⁷Baseline data were missing for 4 patients, 2 siblings and 3 controls. ⁸Baseline data were missing for 150 patients, 115 siblings and 65 controls. ⁹Information was only available for patients. Baseline data were missing for 175 patients. Abbreviations: SD=standard deviation; CPD=number of cigarettes smoked per day; CAPE=Community Assessment of Psychic Experiences.

MacCabe, 2019). A possible explanation of the absence of findings in patients may relate to dominant non-smoking related biological pathways underlying the schizophrenia case-control status. In addition, in patients, intervention changes the symptom course and thereby the association with neurobiological variables (like PRSs for smoking). These associations are likely to be different from siblings and healthy controls who don't receive interventions.

There are contrasting, non-mutually exclusive hypotheses that try to explain the high rates of smoking in patients with psychotic disorders (Quigley & MacCabe, 2019). The shared-vulnerability hypothesis proposes that shared genetic and environmental factors render people more vulnerable to both tobacco use and psychotic disorders (Chambers, Krystal, & Self, 2001). Furthermore, emerging accumulating evidence supports the hypothesis that smoking is a causal risk factor for the development of severe mental illness (Quigley & MacCabe, 2019). For example, a Mendelian randomization study and a meta-analysis of prospective observational studies found evidence for a causal relationship between smoking and schizophrenia (Gurillo, Jauhar, Murray, & MacCabe, 2015; Wootton et al., 2018). However, genetic variation also

captures environmental risk factors. For example, offspring genetic risk for schizophrenia was associated with prenatal environment factors like maternal smoking during pregnancy (Krapohl et al., 2017). Although explained variances were generally small, the results of the current study provide support for a partly shared vulnerability hypothesis of genetic risk for smoking and subclinical levels of psychosis symptoms only in unaffected siblings of schizophrenia patients. Future research is needed to disentangle the complex interplay between smoking and severe mental illness, also by identifying plausible pathophysiological mechanisms via which genetic risks for smoking could affect mental health symptoms (such as the expression of cholinergic receptors) and to elucidate whether this relationship reflects a causal process.

Strengths of our study include the large cohort of patients, siblings and controls, using extensive follow-up phenotype data (enabling multi-cross-sectional associations). Several limitations should nonetheless be noted. First, large samples are required for polygenic risk scores and our study might be underpowered to pick up some small effects in the patient group (Dudbridge, 2013). Second, due to the observational character of our study and the current sample size, causal analyses (e.g.,

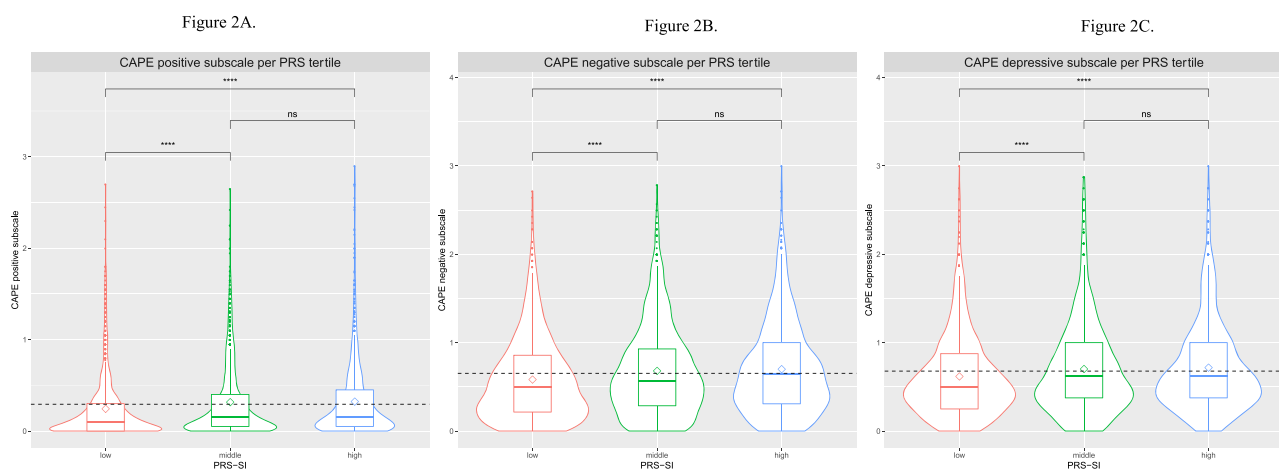


Fig. 2. Violin and box plots of CAPE subscales per PRS-SI tertile for positive symptoms (A), negative symptoms (B) and depressive symptoms (C). Figure 2A. Figure 2B. Figure 2C.

The dashed line denotes the mean CAPE subscale scores in all participants. The diamonds denote the mean CAPE subscales in each group. Mean comparisons were examined using T-tests: ns: $p > 0.05$, *: $p < 0.05$, **: $p < 5.0 \times 10^{-3}$, ****: $p < 5.0 \times 10^{-5}$. As can be appreciated from the graphs, all PRS-SI upper tertiles consistently contained the highest symptom levels relative to the lowest PRS-SI tertiles.

CAPE = community assessment of Psychic Experiences; PRS-SI = polygenic risk scores for smoking initiation.

Figure 3A. Scatter plot of PRS-SI regressed against positive symptoms.

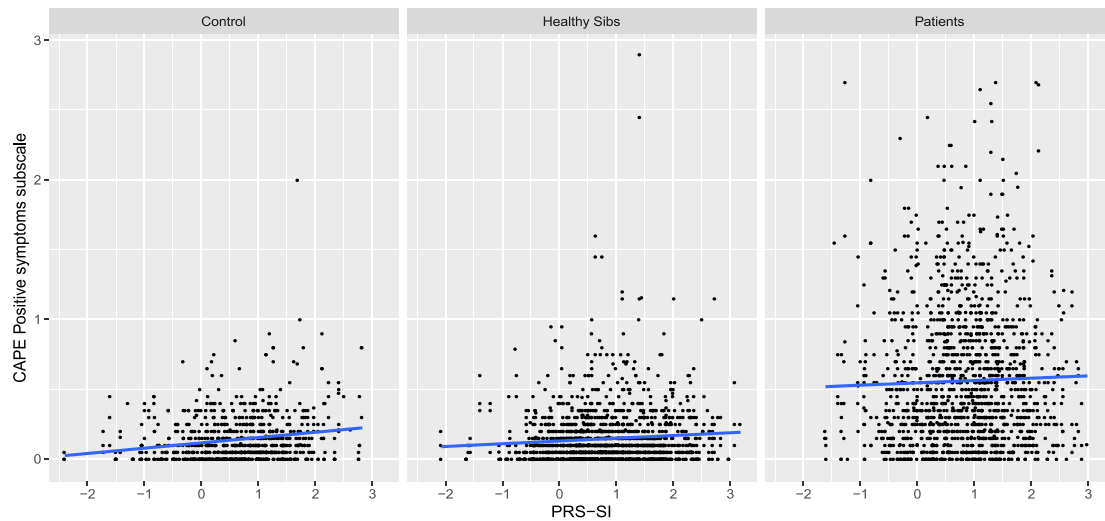


Figure3 B. Scatter plot of PRS-SI and negative symptoms.

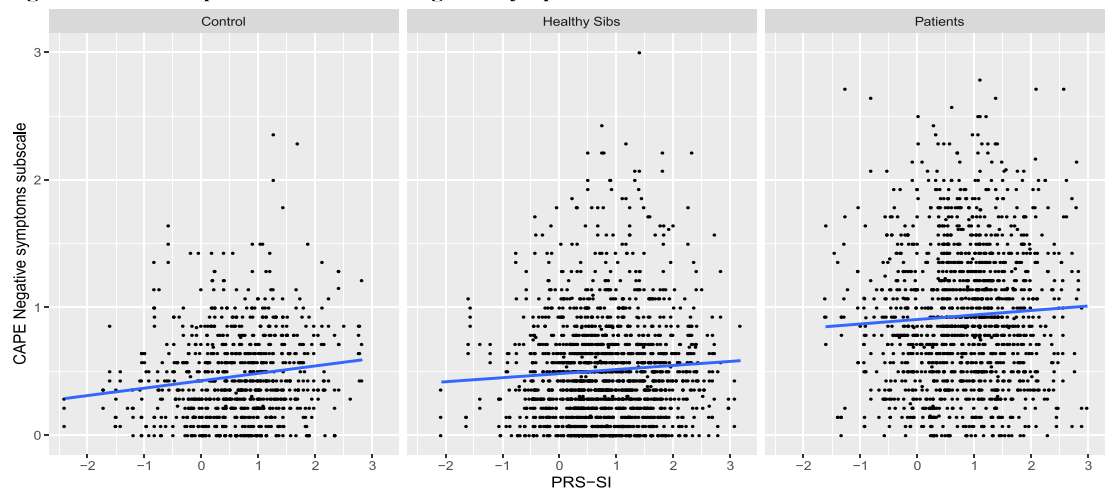


Figure3 C. Scatter plot of PRS-SI and depressive symptoms.

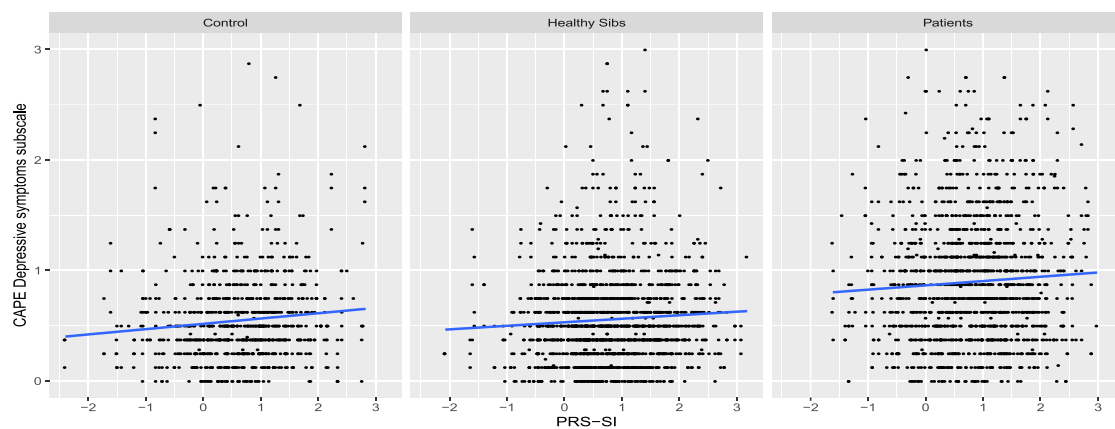


Fig. 3. Scatter plots of polygenic risk scores for smoking initiation (PRS-SI) regressed against psychotic symptoms. **Figure 3A.** Scatter plot of PRS-SI regressed against positive symptoms. **Figure3 B.** Scatter plot of PRS-SI and negative symptoms. **Figure3 C.** Scatter plot of PRS-SI and depressive symptoms. The x-axis represents polygenetic risk scores (PRS), and the y-axis represents symptom scales. The blue line represents the regression line: the slope (effect size of the regression model) decreases from controls to siblings to patients with a psychotic disorder, showing that PRS has larger association results in controls or siblings than in patients. The intercepts increase from controls to siblings to patients, reflecting significant mean symptom score differences between the groups.

one-sample Mendelian randomization) could not be conducted. Future research is needed to explore possible residual confounding effects of environmental factors. Third, patients participating in GROUP reported relatively low levels of psychotic symptoms. This may have influenced our findings and may limit the generalizability of the findings to patients with a more severe course of their psychotic disorder. Finally, some phenotypic substance use data, e.g. the Fagerstrom test for nicotine dependence, have not been collected and may enrich future analyses in other cohorts.

In conclusion, our findings indicate that smoking behaviors in unaffected siblings lie at an intermediate level in between patients and healthy controls. We observed that genetic susceptibility for smoking behavior is associated with schizophrenia symptom expressions in a population of unaffected siblings of patients with a psychotic disorder. Further research into biological pathways is needed to elucidate the underlying mechanisms for this association.

Author statement

JJL and BDL conceived the study. BDL and JMV conducted the analyses. BDL, JMV and JJL wrote the first draft. All authors were involved in data collection and critically revised the final draft and approved of it. JJL supervised the study.

Ethics

This study was approved by the Medical Ethics Committee of the University Medical Center of Utrecht. Written informed consent was obtained before inclusion.

Declaration of Competing Interest

None.

Acknowledgments

We are grateful for the generosity of time and effort by the patients, their families and healthy subjects. Furthermore, we would like to thank all research personnel involved in the GROUP project, in particular: Joyce van Baaren, Erwin Veermans, Ger Driessen, Truda Driessen, Erna van 't Hag.

Financial support

The infrastructure for the GROUP study is funded through the Geestkracht programme of the Dutch Health Research Council (ZonMw, grant number 10-000-1001), and matching funds from participating pharmaceutical companies (Lundbeck, AstraZeneca, Eli Lilly, Janssen Cilag) and universities and mental health care organizations (Amsterdam: Academic Psychiatric Centre of the Academic Medical Center and the mental health institutions: GGZ Ingeest, Arkin, Dijk en Duin, GGZ Rivierduinen, Erasmus Medical Centre, GGZ Noord Holland Noord. Groningen: University Medical Center Groningen and the mental health institutions: Lentis, GGZ Friesland, GGZ Drenthe, Dimence, Mediant, GGNet Warnsveld, Yulius Dordrecht and Parnassia psycho-medical center The Hague. Maastricht: Maastricht University Medical Centre and the mental health institutions: GGzE, GGZ Breburg, GGZ Oost-Brabant, Vincent van Gogh voor Geestelijke Gezondheid, Mondriaan, Virenz riagg, Zuyderland GGZ, MET ggz, Universitair Centrum Sint-Jozef Kortenberg, CAPRI University of Antwerp, PC Ziekeren Sint-Truiden, PZ Sancta Maria Sint-Truiden, GGZ Overpelt, OPZ Rekem. Utrecht: University Medical Center Utrecht and the mental health institutions Altrecht, GGZ Centraal and Delta).

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at [doi:10.1016/j.psychres.2023.115143](https://doi.org/10.1016/j.psychres.2023.115143).

References

- Bates, D., Machler, M., Bolker, B.M., Walker, S.C., 2015. Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* 67 (1), 1–48. Retrieved from <Go to ISI>://WOS:000365981400001.
- Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate - a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B-Stat. Methodol.* 57 (1), 289–300. Retrieved from <Go to ISI>://WOS:A1995QE45300017.
- Chambers, R.A., Krystal, J.H., Self, D.W., 2001. A neurobiological basis for substance abuse comorbidity in schizophrenia. *Biol. Psychiatry* 50 (2), 71–83. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/11526998>.
- Chen, J., Bacanu, S.-A., Yu, H., Zhao, Z., Jia, P., Kendler, K.S., Vanyukov, M., 2016. Genetic relationship between schizophrenia and nicotine dependence. *Sci. Rep.* 6, 25671. <https://doi.org/10.1038/srep25671>.
- Choi, S.W., Mak, T.S., O'Reilly, P.F., 2020. Tutorial: a guide to performing polygenic risk score analyses. *Nat. Protoc.* 15 (9), 2759–2772. <https://doi.org/10.1038/s41596-020-0353-1>.
- Das, S., Forer, L., Schonherr, S., Sidore, C., Locke, A.E., Kwong, A., Fuchsberger, C., 2016. Next-generation genotype imputation service and methods. *Nat. Genet.* 48 (10), 1284–1287. <https://doi.org/10.1038/ng.3656>.
- de Leon, J., Diaz, F.J., 2005. A meta-analysis of worldwide studies demonstrates an association between schizophrenia and tobacco smoking behaviors. *Schizophr. Res.* 76 (2-3), 135–157. <https://doi.org/10.1016/j.schres.2005.02.010>.
- Diagnostic and Statistical Manual of Mental Disorders: DSM-IV-TR.* (2000). Washington: American Psychiatric Publisher.
- Dudbridge, F., 2013. Power and predictive accuracy of polygenic risk scores. *PloS Genet.* 9 (3), e1003348 <https://doi.org/10.1371/journal.pgen.1003348>.
- Gurillo, P., Jauhar, S., Murray, R.M., MacCabe, J.H., 2015. Does tobacco use cause psychosis? Systematic review and meta-analysis. *Lancet Psychiatry* 2 (8), 718–725. [https://doi.org/10.1016/S2215-0366\(15\)00152-2](https://doi.org/10.1016/S2215-0366(15)00152-2).
- Khouja, J.N., Wootton, R.E., Taylor, A.E., Davey Smith, G., Munafo, M.R., 2021. Association of genetic liability to smoking initiation with e-cigarette use in young adults: a cohort study. *PLoS Med.* 18 (3), e1003555 <https://doi.org/10.1371/journal.pmed.1003555>.
- Korver, N., Quee, P.J., Boos, H.B., Simons, C.J., de Haan, L., investigators, G., 2012. Genetic risk and outcome of psychosis (GROUP), a multi-site longitudinal cohort study focused on gene-environment interaction: objectives, sample characteristics, recruitment and assessment methods. *Int. J. Methods Psychiatr. Res.* 21 (3), 205–221. <https://doi.org/10.1002/mpr.1352>.
- Krapohl, E., Hannigan, L.J., Pingault, J.B., Patel, H., Kadeva, N., Curtis, C., Plomin, R., 2017. Widespread covariation of early environmental exposures and trait-associated polygenic variation. *Proc. Natl. Acad. Sci. U. S. A.* 114 (44), 11727–11732. <https://doi.org/10.1073/pnas.1707178114>.
- Kuznetsova, A., Brockhoff, P.B., Christensen, R.H.B., 2017. lmerTest package: tests in linear mixed effects models. *J. Stat. Softw.* 82 (13), 1–26. Retrieved from <Go to ISI>://WOS:000417711600001.
- Lasser, K., Boyd, J.W., Woolhandler, S., Himmelstein, D.U., McCormick, D., Bor, D.H., 2000. Smoking and mental illness: a population-based prevalence study. *JAMA* 284 (20), 2606–2610. <https://doi.org/10.1001/jama.284.20.2606>.
- Liu, M., Jiang, Y., Wedow, R., Li, Y., Brazel, D.M., Chen, F., Psychiatry, H.A.-I., 2019. Association studies of up to 1.2 million individuals yield new insights into the genetic etiology of tobacco and alcohol use. *Nat. Genet.* <https://doi.org/10.1038/s41588-018-0307-5>.
- Lyons, M.J., Bar, J.L., Kremen, W.S., Toomey, R., Eisen, S.A., Goldberg, J., Tsuang, M., 2002. Nicotine and familial vulnerability to schizophrenia: a discordant twin study. *J. Abnorm. Psychol.* 111 (4), 687–693. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/12428784>.
- Mossaheb, N., Becker, J., Schaefer, M.R., Klier, C.M., Schloegelhofer, M., Papageorgiou, K., Amminger, G.P., 2012. The community assessment of psychotic experience (CAPE) questionnaire as a screening-instrument in the detection of individuals at ultra-high risk for psychosis. *Schizophr. Res.* 141 (2-3), 210–214. <https://doi.org/10.1016/j.schres.2012.08.008>.
- Oh, K., Nishizawa, D., Muto, Y., Sugiyama, S., Hasegawa, J., Soda, M., Ikeda, K., 2020. Polygenic risk scores for late smoking initiation associated with the risk of schizophrenia. *NPJ Schizophr.* 6 (1), 36. <https://doi.org/10.1038/s41537-020-00126-z>.
- Olsson, M., Gerhard, T., Huang, C., Crystal, S., Stroup, T.S., 2015. Premature mortality among adults with schizophrenia in the United States. *JAMA psychiatry* 1–10. <https://doi.org/10.1001/jamapsychiatry.2015.1737>.
- Pazoki, R., Lin, B.D., van Eijk, K.R., Schijven, D., de Zwart, S., Investigators, G., Luykx, J.J., 2020. Phenome-wide and genome-wide analyses of quality of life in schizophrenia. *BJPsych Open* 7 (1), e13. <https://doi.org/10.1192/bjo.2020.140>.
- Price, A.L., Weale, M.E., Patterson, N., Myers, S.R., Need, A.C., Shianna, K.V., Reich, D., 2008. Long-range LD can confound genome scans in admixed populations. *Am. J. Hum. Genet.* 83 (1), 132–135. <https://doi.org/10.1016/j.ajhg.2008.06.005> author reply 135–139.
- Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A., Bender, D., Sham, P.C., 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* 81 (3), 559–575. <https://doi.org/10.1086/519795>.

- Quigley, H., MacCabe, J.H., 2019. The relationship between nicotine and psychosis. *Therap. Adv. Psychopharmacol.* 9, 2045125319859969 <https://doi.org/10.1177/2045125319859969>.
- Core Team, R., 2016. R: A Language and Environment for Statistical Computing. Austria: R Foundation for Statistical Computing, Vienna. Retrieved from. <https://www.R-project.org/>.
- Revicki, D., Sobal, J., DeForge, B., 1991. Smoking status and the practice of other unhealthy behaviors. *Fam. Med.* 23 (5), 361–364. Retrieved from. <https://www.ncbi.nlm.nih.gov/pubmed/1884931>.
- Schizophrenia Working Group of the Psychiatric Genomics, C., 2014. Biological insights from 108 schizophrenia-associated genetic loci. *Nature* 511 (7510), 421–427. <https://doi.org/10.1038/nature13595>.
- Trubetskoy, V., Pardinas, A.F., Qi, T., Panagiotaropoulou, G., Awasthi, S., Bigdeli, T.B., 2022. Mapping genomic loci implicates genes and synaptic biology in schizophrenia. *Nature* 604 (7906), 502–508. <https://doi.org/10.1038/s41586-022-04434-5>.
- Vermeulen, J., Schirmbeck, F., Blankers, M., Van Tricht, M., Bruggeman, R., Van den Brink, W., De Haan, L., 2018. Association between smoking behavior and cognitive functioning in patients with psychosis, siblings, and healthy control subjects: results from a prospective 6-year follow-up study. *Am. J. Psychiatry* 0 (0). <https://doi.org/10.1176/appi.ajp.2018.18010069> appi.ajp.2018.18010069.
- Wootton, R. E., Richmond, R. C., Stuijzand, B. G., Lawn, R. B., Sallis, H. M., Taylor, G. M. J., Munafò, M. R. (2018). Causal effects of lifetime smoking on risk for depression and schizophrenia: Evidence from a Mendelian randomisation study. *bioRxiv*. doi: 10.1101/381301.
- Zeng, L.N., Zong, Q.Q., Zhang, L., Feng, Y., Ng, C.H., Ungvari, G.S., Xiang, Y.T., 2020. Worldwide prevalence of smoking cessation in schizophrenia patients: A meta-analysis of comparative and observational studies. *Asian J. Psychiatr.* 54, 102190 <https://doi.org/10.1016/j.ajp.2020.102190>.